UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/768,178	02/02/2004	Xiaohong Huang	122.1580	3651

21171 7590 12/15/2006

STAAS & HALSEY LLP
SUITE 700
1201 NEW YORK AVENUE, N.W.
WASHINGTON, DC 20005

EXAMINER	
DEBROW, JAMES J	

ART UNIT	PAPER NUMBER
2176	

DATE MAILED: 12/15/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/768,178	HUANG ET AL.	
	Examiner	Art Unit	
	James J. Debrow	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *02 February 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-12* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-12* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *02 February 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *02 February 2004*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This action is responsive to communications: Application filled 02 Feb. 2004.

2. Claims 1-12 are pending in this case. Claims 1 and 7 are independent claims.

Specification

3. Applicant is reminded of the proper language and format for an abstract of the disclosure.

The abstract should be in narrative form and generally limited to a single paragraph on a separate sheet within the range of 50 to 150 words. It is important that the abstract not exceed 150 words in length since the space provided for the abstract on the computer tape used by the printer is limited. The form and legal phraseology often used in patent claims, such as "means" and "said," should be avoided. The abstract should describe the disclosure sufficiently to assist readers in deciding whether there is a need for consulting the full patent text for details.

The language should be clear and concise and should not repeat information given in the title. It should avoid using phrases which can be implied, such as, "The disclosure concerns," "The disclosure defined by this invention," "The disclosure describes," etc.

Claim Objections

4. **Claim 2** is objected to because of the following informalities: In line 2, of the claim, the word *special* is mis-spelled. Appropriate correction is required.

5. **Claim 8** objected to because of the following informalities: In line 1 of the claim recites, *"The method according to claim 8...."*. The Examiner believes Applicant meant intended the claim to recite, *"The method according to claim 7...."*, and will be examined as such. Appropriate correction is required.

Claim Rejections - 35 USC § 112

6. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

7. **Regarding claims 1 and 7**, the phrase "such as" renders the claim indefinite because it is unclear whether the limitations following the phrase are part of the claimed invention. See MPEP § 2173.05(d).

8. Regarding **claims 1 and 7**, the term "etc" renders the claim indefinite because it is unclear whether the limitations following the phrase are part of the claimed invention. See MPEP § 2173.05(d).

Claim Rejections - 35 USC § 103

9. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

10. Claims 1-12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Terayama et al. in view of (Pat. No.: 7,010,551 B2; Filed Date: Mar. 15, 2001)

(hereinafter "Terayama"), further in view of Okamoto in view of (Pub. No.: US 2002/0065814 A1; Effective Filing Date: Jun. 30, 1998) (hereinafter "Okamota").

In regards to independent claim 1, Terayama discloses *an apparatus for extracting information from a formatted document, comprising:*

an input unit for inputting a formatted document (col. 2 line 6 – col. 3, line 30; Terayama discloses a file converter for extracting data from a file.);

a unit for identifying special character strings on the basis of the analysis result by means of the typographic information such as font size, character font, color, etc., (col. 1, line 21-29; col. 2 line 6 – col. 3, line 30; Terayama discloses a detector unit for detecting identifiers which indicates the data displayable on a limited-capability device from the file stored in the file storage unit. Terayama also discloses the identifier are called tags which describes *the font, the size, and the color* of the text.).

a unit for extracting the identified special character strings (col. 2 line 6 – col. 3, line 30; Terayama discloses an extractor unit for extracting, from the input file, the data with the start and end of each piece of data. Thus the start and end identifiers (beginning and ending tags) are special character string.).

an output unit for outputting the extracted character strings (col. 2 line 6 – col. 3, line 30; Terayama discloses an output unit for outputting the extracted data.).

Terayama does not disclose expressly *a unit for analyzing the input formatted document and saving the particular typographic information;*

However, Okamota teaches *a unit for analyzing the input formatted document and saving the particular typographic information* (0039-0040; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input (character string / typographic information). Okamota also teaches means for analyzing a structured document input to generate an analyzed structured document and storing the analyzed structured document in a file unit.).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 2, Terayama does not disclose expressly *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

said unit for identifying special character strings determines a certain character string as a special one on the basis of the typographic information of said formatted document when the typographic information of said character string is determined as a special typographic information.

However, Okamota teaches *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

said unit for identifying special character strings determines a certain character string as a special one on the basis of the typographic information of said formatted document when the typographic information of said character string is determined as a special typographic information (0039-0040; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character string / typographic information)*.).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 3, Terayama does not disclose expressly *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

said formatted document is HTML document, and

said unit for identifying special character strings a certain character string as a special one on the basis of the analyzing results with respect to said HTML document

when the font size of said character string is determined to be the biggest one among the surrounding character strings.

However, Okamota teaches *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

said formatted document is HTML document (0026; Okamota teaches searching character strings within a HTML document.).

said unit for identifying special character strings a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings (0037-0040; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to one of ordinary shill in the art that the matching character string *(font size)* being searched as taught by Okamota, could have been *determined to be the biggest one among the surrounding character strings*. Thus Okamota teaches *identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings.* Thus Okamota teaches *identifying special character strings a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the font size*

of said character string is determined to be the biggest one among the surrounding character strings).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 4, Terayama does not disclose expressly *the apparatus for extracting information from a formatted document according to claim 1, wherein*

said formatted document is HTML document,

and said unit for identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings.

However, Okamota teaches *the apparatus for extracting information from a formatted document according to claim 1, wherein*

said formatted document is HTML document (0026; Okamota teaches searching character strings within a HTML document.).

and said unit for identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to one of ordinary shill in the art that the matching character string *(color and the font)* being searched as taught by Okamota, could have been considered *to be a special one among the surrounding character strings.* Thus Okamota teaches *identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings.*).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 5, Terayama does not disclose expressly *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

said formatted document is HTML document, and

said unit for identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the font of said character string is determined to be different from the surrounding character strings and said character string to be boldface.

However, Okamota teaches *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

said formatted document is HTML document (0026; Okamota teaches searching character strings within a HTML document.).

said unit for identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the font of said character string is determined to be different from the surrounding character strings and said character string to be boldface (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to one of ordinary shill in the art that the matching character string *(boldface)* being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings.*).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing

a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 6, Terayama does not disclose expressly *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

said formatted document is HTML document, and

said unit for identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color of said character string is determined to be different from the surrounding character strings and said character string to be boldface.

However, Okamota teaches *the apparatus for extracting information from a formatted document according to claim 1, wherein*

said formatted document is HTML document (0026; Okamota teaches searching character strings within a HTML document.).

said unit for identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color of said character string is determined to be different from the surrounding character strings and said character string to be boldface (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to

one of ordinary shill in the art that the matching character string *(boldface)* being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings.*).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to independent claim 7, Terayama discloses *a method for extracting information from a formatted document, comprising the following steps* (col. 2 line 6 – col. 3, line 30; Terayama discloses a file converter for extracting data from a file.);

identifying special character strings on the basis of the analysis result by means of the typographic information such as font size, character font, color, etc. col. 1, line 21-29; col. 2 line 6 – col. 3, line 30; Terayama discloses a detector unit for detecting identifiers which indicates the data displayable on a limited-capability device from the file stored in the file storage unit. Terayama also discloses the identifier are called tags which describes *the font, the size, and the color* of the text.).

extracting the identified special character strings; and outputting the extracted character strings (col. 2 line 6 – col. 3, line 30; Terayama discloses an extractor unit for extracting, from the input file, the data with the start and end of each piece of data.

Thus the start and end identifiers (beginning and ending tags) are special character string. Terayama further discloses an output unit for outputting the extracted data.).

Terayama does not disclose expressly *inputting a formatted document, analyzing the input formatted document and saving the particular typographic information;*

However, Okamota teaches *inputting a formatted document, analyzing the input formatted document and saving the particular typographic information* (0039-0040; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input (character string / typographic information). Okamota also teaches means for analyzing a structured document input to generate an analyzed structured document and storing the analyzed structured document in a file unit.).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 8, Terayama does not disclose expressly *the method according to claim ~~8~~ 7, wherein in the step of identifying special character string, a certain character string is determined as a special one on the basis of the*

typographic information of said formatted document when the typographic information of said character string is determined as a special typographic information.

However, Okamota teaches *the method according to claim ~~8~~ 7, wherein in the step of identifying special character string, a certain character string is determined as a special one on the basis of the typographic information of said formatted document when the typographic information of said character string is determined as a special typographic information* (0039-0040; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character string / typographic information)*.).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 9, Terayama does not disclose expressly *the method according to claim 7, wherein*

said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the font

size of said character string is determined to be the biggest one among the surrounding character strings.

However, Okamota teaches *the method according to claim 7, wherein said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings*(0037-0040; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to one of ordinary shill in the art that the matching character string *(font size)* being searched as taught by Okamota, could have been *determined to be the biggest one among the surrounding character strings*. Thus Okamota teaches *identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings*. Thus Okamota teaches *identifying special character strings a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings*).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 10, Terayama does not disclose expressly *the method according to claim 7, wherein said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings.*

However, Okamota teaches *the method according to claim 7, wherein said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings* (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to one of ordinary shill in the art that the matching character string *(color and the font)* being searched as taught by Okamota, could have been considered *to be a special one among the surrounding character strings.* Thus

Okamota teaches *identifying special character strings determines a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings.*).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 11, Terayama does not disclose expressly *the method according to claim 7, wherein said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the font of said character string is determined to be different from the surrounding character strings and said character string to be boldface.*

However, Okamota teaches *the method according to claim 7, wherein said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the font of said character string is determined to be different from the surrounding character strings and said*

character string to be boldface (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to one of ordinary shill in the art that the matching character string *(boldface)* being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings.*).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

In regards to dependent claim 12, Terayama does not disclose expressly *the method according to claim 7, wherein said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the color of said character string is determined to be different from the surrounding character strings and said character string to be boldface.*

However, Okamota teaches *the method according to claim 7, wherein said formatted document is HTML document, and in the step of identifying special character string, a certain character string is determined as a special one on the basis of the analyzing results with respect to said HTML document when the color of said character*

string is determined to be different from the surrounding character strings and said character string to be boldface (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input *(character strings)*. It would have been obvious to one of ordinary shill in the art that the matching character string *(boldface)* being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings.*).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Terayama with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

Note

11. It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See, MPEP 2123.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to James J. Debrow whose telephone number is 571-272-5768. The examiner can normally be reached on 8:00-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 571-272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

JAMES DEBROW
EXAMINER
ART UNIT 2176

DOUG HUTTON
PRIMARY EXAMINER
TECH CENTER 2100